Fall 2001

Kernels of Truth or Distorted Perceptions: Self and Observer Ratings of Social Anxiety and Performance

Peter J. Norton
University of Nebraska-Lincoln, peter.norton@monash.edu

Debra A. Hope
University of Nebraska-Lincoln, dhope1@unl.edu

Published in *Behavior Therapy* 32:4 (Autumn 2001), pp. 765–786; doi: 10.1016/S0005-7894(01)80020-4

Submitted June 12, 2000; accepted July 7, 2001; published online March 2, 2006.

Kernels of Truth or Distorted Perceptions: Self and Observer Ratings of Social Anxiety and Performance

Peter J. Norton and Debra A. Hope

University of Nebraska–Lincoln

Corresponding author – Debra A. Hope, Department of Psychology, University of Nebraska–Lincoln, Lincoln, NE 68588-0308, email dhopel@unl.edu

Abstract

This study compared self and observer ratings of social performance and anxiety among individuals with social anxiety disorder, nonclinical controls, and participants with dysthymia serving as clinical controls. The purpose was to elucidate whether self-perceptions of individuals with social anxiety disorder reflect observable performance and anxiety differences, negative self-perceptions, or an interaction of the two. Participants engaged in three role-played interactions, and self and observer ratings of performance and anxiety were obtained. In general, self-ratings of anxiety and performance were more negative (greater anxiety and poorer performance) than were observer ratings. Interactions of rating source and diagnosis indicated the discrepancy between self and observer ratings of both anxiety and performance was significantly greater among participants with social anxiety disorder. Observers, however, generally noted differences across the groups in both anxiety and performance. The discrepancies between self and observer ratings of anxiety were related to negative evaluation fears and negative thought patterns, while performance discrepancies were related to negative thought patterns. Treatment implications are discussed.

Previous research has suggested that impressions of one's own social interactions tend to differ considerably from how others view the interactions. For example, Nelson, Hayes, Felton, and Jarrett (1985) found that college students consistently underrated their social

skill and overrated their anxiety during role-played interactions with opposite-sex confederates, compared to ratings provided by independent observers of the role-plays. These differences between self-ratings and observer ratings appear to be mediated, in part, by social anxiety. Clark and Arkowitz (1975) compared self and observer ratings of social skill and anxiety for high and low socially anxious male college students following a role-play with a female confederate. High-anxiety participants perceived their social skill to be lower than did the observers, while low-anxiety participants perceived their social skill as greater than did the observers. Furthermore, high anxiety participants rated their anxiety higher than did observers, whereas low anxiety participants and observers did not differ on ratings of anxiety. Additional analyses indicated the discrepancies were related to biased self-perceptions, not the groups using different metrics for assessing anxiety and skill.

Along a similar vein, Glasgow and Arkowitz (1975) had high- and low-frequency daters rate their own and their partner's social skill and anxiety following an unstructured 10-minute dyadic interaction. Low-dating-frequency men rated their social skill lower and their anxiety higher than did the high-dating-frequency men, although ratings of skill and anxiety provided by their partner did not differ. Low-dating-frequency women rated their social skill lower than did high-dating-frequency women. However, in contrast to ratings for men, ratings provided by women's partners supported the difference. Interestingly, there were no significant differences between high- and low-dating-frequency women on self or partner ratings of social anxiety.

Extending this line of research to clinical populations, Rapee and Lira (1992) compared the performance of individuals with social anxiety disorder (social phobia) and a control group of nonclinical individuals during a short, unrehearsed speech. After each speech, global ratings of performance and ratings of specific skill performance were assessed by other study participants and the observed participant. Both participants with social anxiety disorder and nonclinical participants rated themselves more poorly on both the specific skills and global impressions than did the observers. Global self-ratings by participants with social anxiety disorder, however, were significantly lower than were global self-ratings made by nonclinical participants, while no difference was found between participants with social anxiety disorder and nonclinical participants on observer ratings of global performance. Thus, although the observers did not detect any difference in performance between participants with social anxiety disorder and nonclinical participants, the results suggest that participants with social anxiety disorder experienced a negative self-perception bias, judging their own performance as poorer than it actually was. Rapee and Lim also noted that, among several measures of social anxiety, self-consciousness, and depression, only fear of negative evaluation was a significant predictor of the discrepancy between self and observer global ratings, accounting for a substantial proportion (31.2%) of the variance in the discrepancy.

Hope, Heimberg, and Bruch (1995) also found that individuals with social anxiety disorder underrated their overall performance during an anxiety-provoking role-play compared to ratings provided by independent observers. Both observer and self-ratings of performance improved following 12 weeks of behavioral or cognitive-behavioral treatment, with the latter increasing to the extent that the difference between observer and self-ratings was no longer significant. No such change was noted for a wait-list control group.

Stopa and Clark (1993) elaborated on the previous studies by employing an anxious, but not socially anxious, control group in comparing self and observer ratings of positive and negative social behaviors displayed during a brief role-played conversation. As expected, participants with social anxiety disorder rated themselves as displaying fewer positive behaviors and more negative behaviors than did observers. Anxious control participants' ratings of their *positive* social behaviors matched the ratings made by observers, but they underrated their own *negative* behaviors relative to the ratings of observers. Interestingly, nonclinical control participants were somewhat more critical of their own performance. Although they did not differ from observers in ratings of negative behaviors, they did report displaying fewer positive behaviors than were reported by the observers. Analysis of observer ratings indicated that socially anxious participants displayed more negative and fewer positive behaviors than either nonclinical or anxious control participants.

Alden and Wallace (1995) assessed self and observer ratings of the degree of visible anxiety exhibited by participants with social anxiety disorder and nonclinical control participants during an unstructured interaction. Although both participants with social anxiety disorder and nonclinical control participants overestimated the visibility of their anxiety in comparison to observer ratings, the degree of overestimation was significantly greater among participants with social anxiety disorder. Furthermore, this interaction of diagnostic group and rating source did not differ between experimental conditions in which the confederate reacted positively or negatively toward the participant. Alden and Wallace, however, did not directly assess differences in observer ratings of anxiety between the diagnostic groups.

Discrepancies between observer and self-ratings of social skill also have been reported in clinical populations other than social anxiety disorder. Lewinsohn, Mischel, Chaplin, and Barton (1980) examined differences in self and observer ratings of a social performance by individuals with depression, an undefined "psychiatric control" group, and a nonclinical control group. Contrary to the findings from socially anxious samples, the results revealed no significant difference in self and observer ratings of performance for participants with depression, but observers rated the performance of nonclinical and clinical control participants significantly lower than their respective self-ratings. Indeed, observers perceived no difference in performance between the three groups, suggesting that the individuals with depression were more accurate in their self-perceptions than were either of the control groups who overestimated the quality of their performance. Interestingly, following a depression treatment program, self-ratings by the participants with depression increased significantly to mirror those of the control groups, although observer ratings did not change significantly following treatment.

Overall, two themes can be drawn from the literature. The first theme relates to actual observed differences in performance and anxiety between socially anxious and nonsocially anxious individuals, and is characterized by mixed conclusions. One study (Stopa & Clark, 1993) suggested observable performance differences between socially anxious and nonsocially anxious individuals, while another (Rapee & Lim, 1992) found no such differences in observer ratings. Interestingly, Glasgow and Arkowitz (1975) reported observed performance differences between high- and low-dating-frequency women, but no differences between high- and low-dating-frequency men.

The second theme is quite consistent across the literature, suggesting that socially anxious individuals tend to perceive their performance in social interactions as more negative than do observers. Although dysfunctional perceptions of one's own social interactions are thought to be a key aspect of social anxiety disorder (Leary & Kowalski, 1995; Rapee & Heimberg, 1997), these may not be unique to social anxiety. Despite this, none of the studies examining perceptions of social functioning among individuals with social anxiety disorder have utilized a nonanxious clinical control group to determine if the social anxiety and poor performance perceived by socially anxious individuals are greater than that expected in a distressed population. In the one study that included an anxious control group, Stopa and Clark (1993) found mixed results when comparing self and observer ratings of positive and negative behaviors between socially anxious participants and the clinical controls. Therefore, this study served to further examine differences between self-ratings and observer ratings of performance and anxiety among participants with social anxiety disorder, in comparison to a nonanxious clinical control group of individuals with dysthymia and matched nonclinical control participants. The purpose was to elucidate whether negative perceptions of one's social performance and anxiety are exaggerated in social anxiety disorder or whether negative perception reflects a common characteristic of individuals experiencing psychological difficulties. The accuracy of those perceptions compared to ratings by objective observers were also explored. Additionally, extending the findings of Rapee and Lim (1992), this study further examined potential variables underlying the discrepancies between self and observer ratings of performance and anxiety.

Individuals with dysthymia were included as a clinical control group for a number of reasons. First, social anxiety disorder and dysthymia have a relatively stable and chronic course, with typical onset in late adolescence or early adulthood (American Psychiatric Association, 1987, 1994). Second, dysfunctional social behavior has been hypothesized to be an important component of both disorders (Becker, 1990; Marks, 1985). Despite these similarities, there is little overlap in the presenting symptoms of the two disorders (American Psychiatric Association, 1987, 1994). The two disorders can co-occur with 11.5% to 31.6% of individuals with social anxiety disorder also meeting lifetime criteria for dysthymia (Wenzel & Holt, 2000).

Several hypotheses were tested. First, it was hypothesized that, consistent with previous research, self-ratings of anxiety would be significantly higher than observer ratings, while self-ratings of performance would be significantly poorer than observer ratings. Second, it was anticipated that the magnitude of the difference between self and observer ratings would vary by diagnosis, wherein participants with social anxiety disorder were expected to show the greatest discrepancy between self and observer ratings of anxiety and performance and nonclinical participants were expected to show the least discrepancy. Third, although the literature is somewhat mixed regarding differences in observed performance and anxiety, it was hypothesized that observers would not detect differences in either performance or anxiety among any of the diagnostic groups. Finally, in light of the Rapee and Lim (1992) data, the differences among the groups in the magnitude of self/observer discrepancy in anxiety and performance ratings were expected to be related to fears of negative evaluation.

Methods

Participants

Participants were recruited as part of a larger study on anxiety and depression via flyers, radio, newspaper, and television advertisements and public service announcements. Separate advertisements invited individuals with social anxiety or depression to receive no-cost treatment in exchange for research participation. The flyers and print advertisements invited persons without psychological difficulties to participate in the research for a financial payment. All potential participants were assessed using the Anxiety Disorders Interview Schedule-Revised (ADIS-R; DiNardo & Barlow, 1988) and the depression section of the Structured Clinical Interview for *DSM-III-R* (SCID; Spitzer, Williams, Gibbon, & First, 1989). Data collection began prior to the release of *DSM-IV* (American Psychiatric Association, 1994), so diagnoses were all based on *DSM-III-R* (APA, 1987) criteria. There was little change in the criteria for Social Anxiety Disorder and Dysthymia, however, from *DSM-III-R* and *DSM-IV*, and we are thus confident that the results are applicable within our current nosological system.

Advanced doctoral students and a licensed clinical psychologist conducted the diagnostic interviews, and had met rigorous training standards for reliability with an expert ADIS-R interviewer. Training included observing three interviews conducted by an experienced interviewer, then conducting at least five interviews under observation, matching the experienced interviewer on four of five diagnoses and matching the Clinician Severity Rating (CSR; see below) within 1 point for the principal diagnosis. All cases were presented at staff meetings and diagnoses were reviewed until a consensus was achieved. Any diagnostic questions were resolved by an additional telephone interview until a consensus diagnosis was reached. Finally, the second author supervised treatment of all of the cases. Despite this extensive contact, no diagnostic errors appeared.

Individuals who met *DSM-III-R* criteria for a principal diagnosis of social anxiety disorder (social phobia; see Liebowitz, Heimberg, Travers, & Stein, 2000) or dysthymia, and had a ADIS-R CSR of 4 (indicating moderate severity) or greater on the 0-to-8 scale were invited to participate. Individuals with comorbid anxiety and depression diagnoses were accepted into the study as long as social anxiety disorder or dysthymia was determined to be their principal diagnosis, as indicated by a higher CSR rating, and they did not meet exclusion criteria. Individuals with comorbid conditions were included in this study in an effort to make the sample more representative of treatment-seeking samples. Indeed, recent studies suggest as high as 50% to 60% comorbidity within and across the anxiety and mood disorders (Brown & Barlow, 1992; Sanderson, DiNardo, Rapee, & Barlow, 1990).

Exclusion criteria were current drug or alcohol abuse or dependence, psychotic or thought disorder, developmental disability, or suicidality requiring immediate intervention. Clinical participants were offered no-cost treatment in exchange for participation in the larger psychopathology study. Individuals who either did not meet criteria for any disorder or met criteria only for specific phobia of subclinical severity (CSR < 4), passed exclusion criteria, and had no history of receiving mental health services, were included in the study as nonclinical control participants. Nonclinical participants were paid $50 for

participation in the overall study. Additional details regarding subject recruitment and selection are described elsewhere (Waiters & Hope, 1998; Weilage & Hope, 1999).

Following screening, 105 participants met the standards for inclusion in the study. Participants were 54 (60.4% women) individuals meeting *DSM-III-R* criteria for social anxiety disorder, 23 (82.6% women) meeting *DSM-III-R* criteria for dysthymia, and 28 (64.3% women) who did not meet criteria for any Axis I diagnosis. Furthermore, of those diagnosed with social anxiety disorder, 32 (59.3%) met criteria for the generalized and 22 met criteria for the nongeneralized subtype as defined by Heimberg, Holt, Schneier, Spitzer, and Liebowitz (1993). Of those diagnosed as having social anxiety disorder, 17 (31.5%) had a secondary diagnosis of a mood disorder with an average CSR of 3.53, with 2 at a level of clinical severity (CSR ≥ 4). Secondary comorbid Axis I disorders included panic disorder with and without agoraphobia (n = 2), obsessive-compulsive disorder (n = 2), specific phobias (n = 20), generalized anxiety disorder (n = 9), unspecified adjustment disorder (n = 1), and hypochondriasis (n = 1). Axis II disorders were not assessed. Among those diagnosed with dysthymia, 9 (39.1%) were given a secondary diagnosis of social anxiety disorder with an average CSR of 2.78, and none of clinical severity (CSR ≥ 4). Other secondary Axis I disorders included major depressive disorder (n = 8), panic disorder with or without agoraphobia (n = 4), obsessive-compulsive disorder (n = 1), generalized anxiety disorder (n = 3), and specific phobias (n = 9). Again, Axis II disorders were not assessed.

Self-Report Questionnaires

For the larger study, participants completed a battery of self-report measures including the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) and the Brief Fear of Negative Evaluation Scale (BFNE; Leary, 1983). The BDI is an extensively used 21-item measure of the somatic, cognitive, and affective domains of depression. The original BDI was used as data collection commenced before publication of the updated version. Beck, Steer, and Garbin (1988) report that the BDI has demonstrated excellent reliability and validity in use with clinical and nonclinical populations. The BFNE, an adaptation of the Fear of Negative Evaluation Scale (FNES; Watson & Friend, 1969), is a 12-item scale assessing fear of negative evaluation stemming from perceived loss of social approval. The BFNE correlates nearly perfectly to the original FNES (r = .96), and both have demonstrated adequate reliability and validity (Leary).

Behavioral Tests

Each participant engaged in three role-played scenarios: (a) a brief speech, (b) an unstructured role-play, and (c) a structured role-play. The order of the scenarios was randomized across participants.

Speech

Participants were given 3 minutes to prepare a 4-minute speech on either a personal interest or hobby, a vacation they had taken, or their occupation. Following the 3-minute preparation, the participant delivered the speech to an audience consisting of the experimenter and two research assistants. The speech format was used because it is feared by nearly all individuals with social anxiety disorder (Holt, Heimberg, Hope, & Liebowitz, 1992), and

it generally elicits an anxiety response under laboratory conditions (Beidel, Turner, Jacob, & Cooley, 1989).

Unstructured role-play

The unstructured role-play involved a 4-minute conversation with a research assistant. The participant and research assistant were instructed to get to know each other better, and could discuss any topic except the study. The confederates were not given specific training except to act as they would in a nonlaboratory situation, speaking as much or as little as seems appropriate. All potential confederates, however, engaged in a similar role-play with a graduate assistant. Those who dominated the conversation, were underinvolved in the conversation, or had other unusual interpersonal behavior, were not used in the role-plays with participants. The unstructured role-play was included in the design, as it allows the role-play to evolve naturally with the confederate varying her behavior depending upon what the participant does. To maximize the natural flow of the interaction, no confederate completed more than 10 role-plays. Similar unstructured interactions are commonly used in research on interpersonal behavior in depression (e.g., Lewinsohn et al., 1980).

Structured role-play

The structured role-play involved a 4-minute conversation with a confederate who had been trained to act in a friendly but reserved manner. The scenario was meeting a new neighbor who had just moved in. Such role-plays are commonly used in studies of social anxiety disorder (e.g., Heimberg, Dodge, Hope, Kennedy, & Zollo, 1990; Hope, Herbert, & White, 1995) as the structure of the role-play, including the confederate's behavior, provides a standard stimulus against which to evaluate the participant's behavior.

Anxiety and performance ratings

Immediately following each role-play, participants rated their peak anxiety during the role-play and their perception of the quality of their performance on 0 to 100 scales used in previous research (e.g., Heimberg, Dodge, et al., 1990; Hope, Heimberg, et al., 1995). The anxiety scale had descriptors at each quartile and the performance scale had descriptors at 0, 50, and 100. Performance ratings were then reversed to aid discernment of similarities and differences between performance and anxiety ratings. Higher numbers indicate greater anxiety and poorer performance quality.

At the completion of the data collection phase of the larger study, three undergraduate research assistants rated videotapes of the three behavioral tests using the same 0-to-100 peak anxiety and performance ratings. Performance ratings were again reversed. The raters were unaware of participant diagnosis, participant self-ratings of anxiety and performance, and the hypotheses of the study. As in previous research (Hope, Heimberg, et al., 1995), the goal was for the raters to provide a consensus opinion as to how participants might be viewed if their performance in the behavioral tests had occurred in their daily lives. Thus, raters were not specifically trained or given instructions about the ratings. All raters did view tapes of two actual participants prior to beginning the ratings. Two tapes, one in which

the participant was extremely anxious and another in which the participant was quite comfortable, were selected to provide some general sense of potential endpoints of the scale. However, this was not explicitly discussed with the raters and was done in the context of teaching them to use the video equipment and explaining the procedures. Each rater viewed the randomly ordered tapes independently.

The interobserver reliabilities of the performance and anxiety ratings provided by the three observers were analyzed using the Spearman-Brown formula. Overall, interobserver reliabilities were acceptable, ranging from .72 to .79 (M = .75) for anxiety ratings and from .74 to .84 (M = .80) for performance ratings. For analyses, the ratings were averaged, yielding three mean anxiety and performance ratings for each participant across the three behavioral tests.

Thought listing

Following each role-play, participants were asked to list all of the thoughts they recalled experiencing during the role-play on a prepared form (Cacioppo, Glass, & Merluzzi, 1979). Thoughts were later coded by research assistants, who were unaware of participant diagnosis, into positive, neutral, or negative thoughts. The raters were highly reliable in coding thoughts (κ = .93). As recommended by Heimberg, Bruch, Hope, and Dombeck, (1990), the total number of negative thoughts was summed and divided by the overall number of thoughts, resulting in a new variable reflecting the percent of the total number of thoughts that were negative.

Confederates' affect ratings

Before and after each unstructured role-play, the confederate completed the state form of the Multiple Affect Adjective Checklist-Revised (MAACL-R; Zuckerman & Lubin, 1985) to assess the impact of the participant on the confederate. The MAACL-R consists of 132 adjectives to describe one's affective state. Adjectives are clustered into five subscales that assess anxiety, depression, hostility, positive affect, and sensation seeking. Lubin, Zuckerman, Hanson, Armstrong, and Rinck (1986) report that the MAACL-R has adequate psychometric properties. Due to the potential error associated with change scores (cf. Keppel & Zedeck, 1989), each subscale score obtained after the unstructured role-play was regressed onto the respective subscale score provided before the unstructured role-play. Standardized residual scores were retained and used as a measure of the impact of the unstructured role-play on the confederate. The MAACL-R was collected only for the unstructured role-play because it most approximated a naturalistic conversation. We were interested in the affective response the socially anxious and dysthymic participants might generate in others during social interactions.

Confederates

All confederates were undergraduate research assistants. Confederates in the structured and unstructured role-plays were Caucasian women, but both men and women served as audience members for the speech. Only women were recruited as confederates for the interactions in order to limit variability due to gender of the confederate. Given that some

participants were older, married, or gay/lesbian, opposite-sex confederates would not necessarily have been potential dating partners, making the nature of the interactions difficult to interpret.

All confederates were unaware of the hypotheses and participants' diagnoses. As noted above, confederates for the structured role-play were trained to respond in a neutral but friendly manner, letting the participant carry the conversation. The experimenter monitored the role-plays and gave corrective feedback as needed. Confederates for the unstructured conversation were instructed to interact as they would if they encountered this person outside of the laboratory. Audience members for the speech were instructed to maintain neutral but attentive facial expressions. Some of the confederates for the unstructured role-play were later trained as confederates for the structured role-play for different participants. Once trained, confederates were not allowed to serve in the unstructured role-play again.

Procedure

Following a brief telephone screening, potential participants were scheduled for the ADIS-R interview. Those who met criteria for the study were given a packet of self-report questionnaires to complete at home and return at the subsequent appointment. Participants were scheduled for the behavior tests at a second appointment within the week. During the second appointment, they also completed some information processing measures not included in the current study. Participants who met criteria for social anxiety disorder or dysthymia were then included in the next available cognitive-behavioral treatment group.

Results

For all analyses, alpha was set at .05. Significant main effects involving more than two groups (i.e., diagnostic group or behavioral test situation) were subsequently followed up using unprotected (LSD) post-hoc analyses ($\alpha = .05$).

Preliminary Analyses

Summary demographic data for each of the diagnostic groups are presented in Table 1. No significant differences were found across the diagnostic groups in terms of age, gender, marital status, or education. Nonclinical control participants and participants with social anxiety disorder, however, were more likely to be working full-time than participants with dysthymia. Furthermore, Multivariate Analysis of Variance (MANOVA) indicated the multivariate main effect for gender on self and observer ratings of anxiety, $F(6, 93) = 1.51$, $p = 19$, Pillai = .09, or performance, $F(6, 92) = 1.79$, $p = .11$, Pillai = .11, was not significant.

One-way analyses of variance (ANOVA) were conducted to confirm the distinctiveness of the diagnostic groups. As expected, participants with dysthymia ($M = 23.96$, $SD = 11.24$) scored significantly higher on the BDI than did participants with social anxiety disorder ($M = 16.23$, $SD = 9.21$) who, in turn, scored significantly higher than did nonclinical control participants ($M = 2.39$, $SD = 2.53$), $F(2,101) = 43.27$, $p < .001$, $d = 1.31$. Participants with social anxiety disorder ($M = 48.21$, $SD = 7.86$) reported significantly higher BFNE scores than did

participants with dysthymia (M = 40.86, SD = 10.86) and nonclinical controls (M = 24.00, SD = 5.18), who differed significantly as well, $F(2, 90) = 77.75$, $p < .001$, $d = 1.86$.

Table 1. Demographic Information by Diagnostic Group

	Social Anxiety Disorder	Dysthymia	Nonclinical	Comparison
Gender				
Men	22 (39.6%)	4 (17.4%)	10 (35.7%)	$\chi^2(2) = 3.62$
Women	32 (60.4%)	19 (82.6%)	18 (64.3%)	$p = .16$
Age	$M = 39.3$	$M = 41.7$	$M = 37.4$	$F(2, 101) = 1.05$
	$SD = 10.1$	$SD = 10.0$	$SD = 11.9$	$p = .37$
Marital Status[a]				
Married	15 (28.3%)	10 (43.5%)	17 (60.7%)	$\chi^2(2) = 4.35$
Single	25 (47.2%)	6 (26.1%)	8 (28.6%)	$p = .11$
Divorced	7 (13.2%)	6 (26.1%)	3 (10.7%)	
Separated	3 (5.7%)	1 (4.3%)	0 (0.0%)	
Widow(er)	3 (5.7%)	0 (0.0%)	0 (0.0%)	
Employment Status[b]				
Full-time	40 (75.5%)	13 (56.5%)	25 (89.3%)	$\chi^2(2) = 7.24$
Part-time	3 (5.7%)	3 (13.0%)	0 (0.0%)	$p = .03$
Student	5 (9.4%)	4 (17.4%)	3 (10.7%)	
Homemaker	1 (1.9%)	1 (4.3%)	0 (0.0%)	
Unemployed	3 (5.7%)	2 (18.7%)	0 (0.0%)	
Retired	1 (1.9%)	0 (0.0%)	0 (0.0%)	
Highest Education				
High school	10 (18.9%)	5 (21.7%)	4 (14.3%)	$\chi^2(4) = 2.08$
Undergraduate	36 (67.9%)	14 (60.9%)	17 (60.7%)	$p = .72$
Graduate school	7 (13.2%)	4 (17.4%)	7 (25.0%)	

a. Due to empty cells, χ^2 compared single to those who are or have been married.
b. Due to empty cells, χ^2 compared full-time to all other employment statuses.

The impact of comorbid social anxiety and mood disorders was examined for the primary dependent variables. Participants with social anxiety disorder who were or were not given a secondary mood disorder diagnosis were compared on self and observer ratings of anxiety and performance for each of the behavioral tests, and no between group differences were found [anxiety: $F(6, 43) = 1.24$, $p = .31$, Pillai = .15; performance: $F(6, 43) = 1.23$, $p = .31$, Pillai = .15]. Similar analyses were conducted for participants with dysthymia who were or were not given a secondary diagnosis of social anxiety disorder; again, no between group differences on self and observer ratings were found [anxiety: $F(6, 16) = 1.83$, $p = .16$, Pillai = .41; performance: $F(6, 15) = 0.67$, $p = .68$, Pillai = .21]. As these MANOVAs may have been underpowered because of limited sample size, t tests were conducted to compare the noncomorbid and comorbid subgroups on computed measures of the discrepancies between self and observer ratings of performance and anxiety (described below). Converging with the MANOVA results, no differences were noted between participants with social anxiety disorder who did or did not have a comorbid mood disorder [anxiety: $t(48) = 0.24$, $p = .81$; performance: $t(48) = 1.72$, $p = .09$], nor between participants with dysthymia who did or did not have comorbid social anxiety disorder [anxiety: $t(21) = 0.90$, $p = .38$; performance: $t(20) = 0.51$, $p = .62$].

Self and Observer Ratings of Anxiety

To analyze differences in anxiety ratings, a 2 (Rating Source: Self vs. Other) × 3 (Assessment Situation: Speech vs. Unstructured Role-Play vs. Structured Role-Play) × 3 (Diagnostic Group: Social Anxiety Disorder vs. Dysthymia vs. Normal Control) ANOVA with repeated measures on the first two factors was performed with anxiety ratings as the dependent variable. Univariate summary statistics of anxiety ratings are provided in Table 2.

Significant main effects of diagnosis, $F(2, 97) = 35.61$, $p < .001$, $d = 1.21$, rating source, $F(1,97) = 52.59$, $p < .001$, $d = 1.47$, and assessment situation $F(2, 96) = 8.68$, $p < .001$, $d = .60$, were found. The main effect of diagnosis was such that nonclinical participants demonstrated lower anxiety than participants with dysthymia who had lower ratings than did participants with social anxiety disorder. The main effect of rating source indicated that, as hypothesized, self-ratings of anxiety were significantly higher than were ratings made by the observers. Finally, the main effect of situation was such that lower anxiety was reported during the unstructured role-play than during either the structured role-play or the speech, which did not differ.

Table 2. Means and Standard Deviations for Anxiety Ratings Made by Observers and Participants

	Observer Ratings			Self-Ratings		
	Speech	Unstructured Role-Play	Structured Role-Play	Speech	Unstructured Role-Play	Structured Role-Play
Social Anxiety Disorder	40.42 (12.77)	39.83 (10.72)	42.33 (12.37)	73.60 (19.87)	66.10 (19.09)	76.20 (17.94)
Dysthymia	30.94 (9.67)	35.98 (11.73)	37.47 (12.50)	58.91 (28.52)	50.43 (23.50)	57.09 (24.51)
Non-Clinical	29.64 (10.65)	30.33 (12.63)	32.43 (10.57)	35.30 (22.52)	28.70 (22.30)	31.48 (23.28)

Note: $n = 50$ for social anxiety disorder; $n = 23$ for dysthymia; $n = 27$ for nonclinical participants. Standard deviations are in parentheses. Ratings made on a 0-to-100 scale, with higher scores indicating greater anxiety.

Analysis of two-way effects revealed no significant interaction of assessment situation and diagnosis, $F(4, 194) = 0.58$, $p = .68$, $d = .11$. A significant interaction of rating source and assessment situation was found, $F(2, 96) = 9.64$, $p < .001$, $d = .63$. The interaction was such that, on observer-ratings, anxiety was significantly higher during structured role-plays than either unstructured role-plays or speeches, which did not differ. For self-ratings, however, anxiety was significantly higher during speeches and structured role-plays, which did not differ, than during unstructured role-plays.

As expected, an interaction of rating source and diagnosis was found, $F(2, 97) = 22.94$, $p < .001$, $d = .97$ (see Fig. 1). To elucidate the nature of the interaction and to test the second hypothesis, a variable reflecting the discrepancy between self and observer anxiety ratings was computed. Due to problems in partitioning variance in simple difference scores (cf. Keppel & Zedeck, 1989), the average self-rating of anxiety across situations was regressed onto the average observer rating of anxiety across situations and the standardized residual score was retained as a measure of discrepancy. The discrepancy score was then included as the

dependent measure in a one-way ANOVA, with diagnostic group as the independent factor, and LSD post-hoc analyses were conducted. Supporting the hypothesis, post-hoc analyses indicated that participants with social anxiety disorder had a greater self vs. observer anxiety discrepancy than either participants with dysthymia or nonclinical participants and participants with dysthymia showed a trend (p = .08) toward greater self vs. observer discrepancy than did nonclinical participants.

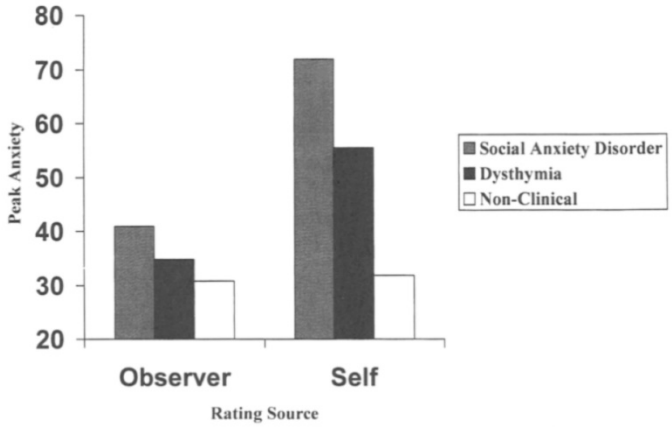

Figure 1. Self and observer anxiety ratings for participants with social anxiety disorder or dysthymia and nonclinical controls.

Testing part of the third hypothesis, observer anxiety ratings were averaged across the three situations and included in a one-way ANOVA to compare ratings across diagnostic groups. Partially refuting the third hypothesis, the results indicated that observers rated participants with social anxiety disorder as significantly more anxious than either nonclinical participants or participants with dysthymia, but the latter two groups did not differ significantly from each other.

Examination of the three-way effect revealed no significant interaction of rating source, assessment situation, and diagnostic group, $F(4, 194) = 1.59$, $p = .18$, $d = .18$, indicating that the above-reported interaction of diagnosis and rating source was consistent across each of the three assessment situations.

Due to concerns that the cross-comorbidity between the two clinical groups may have influenced the results, the 2 × 3 × 3 ANOVA was reanalyzed twice. First, the ANOVA was reanalyzed after removing the data from both participants with social anxiety disorder who had a comorbid mood disorder, and participants with dysthymia who had comorbid social anxiety disorder. Second, the ANOVA was reanalyzed after removing data from individuals whose comorbid mood disorder or social anxiety disorder met criteria for clinical severity (CSR ≥ 4). Effect sizes from each of the main effects and interaction effects were then converted to the *Z* statistic and compared using the Fisher *Z*-test. None of the main effects or interaction effects from the two reduced ANOVAs differed significantly from the corresponding main effects and interaction effects from the full ANOVA (all *Z*s < .65, *ns*).

Self and Observer Ratings of Performance

Ratings of participant performance were analyzed using a similar 2 (Rating Source) × 3 (Assessment Situation) × 3 (Diagnostic Group) ANOVA with repeated measures on the first two factors. Univariate summary statistics of performance ratings are provided in Table 3.

Table 3. Means and Standard Deviations for Performance Ratings Made by Observers and Participants

	Observer Ratings			Self-Ratings		
	Speech	Unstructured Role-Play	Structured Role-Play	Speech	Unstructured Role-Play	Structured Role-Play
Social Anxiety Disorder	37.35 (12.19)	36.13 (9.29)	39.60 (20.74)	60.60 (20.74)	54.80 (18.98)	66.20 (18.83)
Dysthymia	31.25 (8.97)	29.66 (10.36)	49.55 (24.97)	49.55 (24.97)	42.50 (25.15)	52.95 (26.40)
Non-Clinical	23.09 (10.85)	24.53 (10.51)	37.78 (15.34)	37.78 (15.34)	34.63 (18.50)	36.85 (20.44)

Note: n = 50 for social anxiety disorder; n = 22 for dysthymia; n = 27 for nonclinical participants. Standard deviations are in parentheses. Ratings made on a 0-to-100 scale, with higher scores indicating greater anxiety.

Significant main effects of diagnosis, $F(2, 96) = 22.03$, $p < .001$, $d = .96$, rating source, $F(1,96) = 90.12$, $p < .001$, $d = 1.94$, and assessment situation, $F(2, 95) = 16.36$, $p < .001$, $d = .83$, were found. Examination of the main effect of diagnosis revealed that nonclinical participants performed significantly better than did participants with either dysthymia or social anxiety disorder, and participants with dysthymia performed significantly better than did participants with social anxiety disorder. The main effect of rating source was such that, as hypothesized, observers rated performance as significantly superior than did the participants themselves. Finally, the main effect of assessment situation was such that performance during the unstructured role-play was significantly better than performance during either the structured role-play or the speech, and performance during the speech was significantly better than performance during the structured role-play.

Analysis of two-way interactions revealed no significant interaction of diagnosis and assessment situation, $F(4, 192) = 1.40$, $p = .23$, $d = .17$. There was, however, a significant interaction of rating source and assessment situation, $F(2, 95) = 3.25$, $p = .04$, $d = .37$. The pattern of the interaction was such that for observer ratings, performance was significantly worse during the structured role-play than during either the speech or unstructured role-play, which did not differ significantly from each other. For self-ratings, performance was significantly superior during the unstructured role-play than during either the speech or structured role-play, which did not differ significantly.

As expected, there was a significant interaction effect of rating source and diagnosis, $F(2, 96) = 4.17$, $p = .02$, $d = .42$ (see Figure 2). To examine the interaction and to further test the second hypothesis, a variable reflecting the discrepancy between self and observer performance ratings was computed. As with the anxiety ratings, the average self-rating of performance across situations was regressed onto the average observer rating of performance across situations and the standardized residual score was retained as a measure of discrepancy. The discrepancy score was then included as the dependent measure in a one-

way ANOVA with diagnostic group as the independent factor, and LSD post-hoc analyses were conducted. Only partially supporting the hypothesis, the post-hoc analyses indicated that participants with social anxiety disorder had a greater self vs. observer performance discrepancy than nonclinical participants. Participants with dysthymia did not differ on performance discrepancy scores from either participants with social anxiety disorder or nonclinical participants.

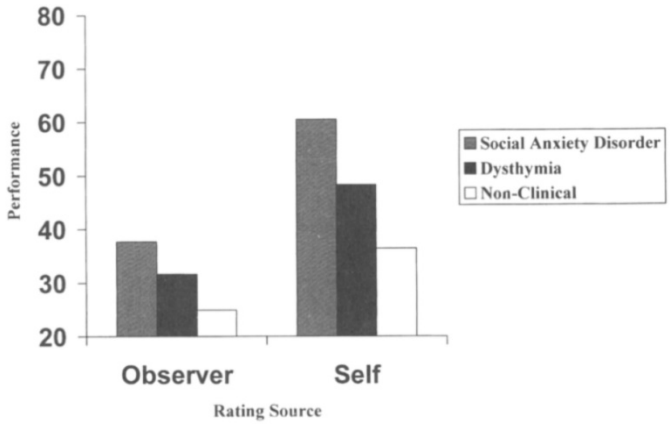

Figure 2. Self and observer performance ratings for participants with social anxiety disorder or dysthymia and nonclinical control.

To further test the third hypothesis, that observers would not detect differences in performance among the three groups, observer performance ratings were averaged across the three situations and included in a one-way ANOVA to compare ratings across diagnostic groups. Contrary to the hypothesis, the results indicated that observers' ratings differed significantly across the three groups, with the performance of individuals with social anxiety disorder rated poorest and the performance of the nonclinical participants rated as the best. Examination of the three-way effect again indicated no significant interaction of rating source, assessment situation, and diagnostic group, $F(4, 192) = 1.19$, $p = .32$, $d = .16$, indicating that the above-reported interaction of diagnosis and rating source was consistent across each of the three assessment situations.

As was described previously for the anxiety ratings, the 2 × 3 × 3 performance rating AVOVA was reanalyzed twice to examine if cross-comorbidity may have impacted the preceding data. None of the main effects or interaction effects from the two reduced ANOVAs differed significantly from the corresponding main effects and interaction effects from the full ANOVA (all Zs < .91, *ns*).

Predictors of Discrepancies

To explicate the nature of the self/observer discrepancies in anxiety ratings and performance ratings and test the final hypothesis, the previously computed anxiety and performance residual scores were used as criterion variables in separate stepwise multiple regression analyses employing the percent of negative thoughts, BDI and BFNE scores,

and confederate MAACL-R subscale residual scores as predictor variables. Stepwise regression was selected to maximize the comparability between our results and those of Rapee and Lira (1992). The resulting regression equation employing anxiety discrepancy scores as the criterion indicated a significant linear relationship, $R^2 = .41$, $F(2, 86) = 29.63$, $p < .001$, $d = 1.67$. Only BFNE, $r_{partial} = .50$, $\beta = .48$, $t = 5.06$, $p < .001$, $d = 1.15$, and percent negative thoughts, $r_{partial} = .28$, $\beta = .26$, $t = 2.71$, $p = .01$, $d = .58$, were significant unique contributors to the prediction of the anxiety rating discrepancy. BDI, $t = 0.48$, $p = .63$, $d = .10$, and confederate MAACL-R subscale residual scores for anxiety, $t = 0.66$, $p = .51$, $d = .14$, depression, $t = 0.14$, $p = .89$, $d = .03$, hostility, $t = 1.15$, $p = .25$, $d = .25$, positive affect, $t = 0.36$, $p = .72$, $d = .08$, and sensation seeking, $t = 0.54$, $p = .59$, $d = .12$, failed to make significant, unique contributions.

Performance discrepancy scores were used as the criterion in a second stepwise regression using the same set of predictors. The final regression equation was significant, $R^2 = .23$, $F(1, 86) = 25.87$, $p < .01$, $d = 1.09$, with only percent negative thoughts making a significant unique contribution, $r_{partial} = .48$, $\beta = –.48$, $t = 5.09$, $p < .001$, $d = 1.09$. Neither BFNE, $t = 1.36$, $p = .18$, $d = .30$, BDI, $t = 0.75$, $p = .46$, $d = .16$, nor confederate MAACL-R anxiety residual score, $t = 0.60$, $p = .55$, $d = .13$, depression residual score, $t = 0.09$, $p = .93$, $d = .02$, hostility residual score, $t = 0.45$, $p = .65$, $d = .10$, positive affect residual score, $t = 0.37$, $p = .71$, $d = .08$, or sensation seeking residual score, $t = 0.25$, $p = .81$, $d = .05$, made significant unique contributions.

Discussion

Consistent with the first hypothesis and much of the research comparing self and observer ratings of social anxiety and social performance (e.g., Hope, Heimberg, et al., 1995; Nelson et al., 1985; Rapee & Lim, 1992; but see Clark & Arkowitz, 1975), psychologically distressed participants in this study demonstrated a negative bias in their self-ratings, reporting their anxiety as higher and performance as poorer than did independent observers. This result, along with the previous research, may suggest that individuals experiencing psychological difficulties tend to perceive themselves more negatively than do others. The degree of self-depreciation exhibited, however, and the relationship to diagnosis, varies when self-evaluating anxiety and social performance.

Partial support was found for the second hypothesis that the magnitude of the difference between self and observer ratings would vary by diagnosis. The discrepancy between self and observer ratings of anxiety conformed to the hypothesized pattern, wherein the discrepancy was greater for participants with social anxiety disorder than participants with dysthymia and nonclinical participants, and greater for participants with dysthymia than nonclinical participants. Participants with social anxiety disorder also showed a greater self-observer discrepancy in performance ratings than did nonclinical participants, although, somewhat discrepant to the second hypothesis, participants with dysthymia did not differ from either group.

The third hypothesis, that observers would not detect differences in anxiety or performance among the groups, received only minimal support. The observers' anxiety and performance ratings differed among the diagnostic groups with the exception that the observer

ratings of anxiety did not differ between participants with dysthymia and nonclinical participants.

In testing the final hypothesis, fear of negative evaluation and percent of thoughts that were negative were found to be unique predictors of the self/observer discrepancy in anxiety ratings. Furthermore, only percent of negative thoughts was uniquely related to the self/observer discrepancy in performance ratings. The lack of relationship between fear of negative evaluation and discrepancy in performance ratings does not replicate the findings of Rapee and Lim (1992), who found fear of negative evaluation to be related to a greater discrepancy between self and observer global ratings of performance quality during a social interaction. It is possible that the difference between the current findings and those of Rapee and Lim is due to patterns of collinearity between the sets of predictors used in the two studies. Although assessing similar constructs, the two studies employed different sets of self-report measures as predictors which, due to the various intercorrelations, could have altered the extent to which any single variable made a unique contribution to the regression. Indeed, BFNE alone significantly predicted performance discrepancy scores, $r^2 = .12$, $p < .05$, although it accounted for less variance than was found by Rapee and Lim, $r^2 = .31$. It is also possible that the use of separate anxiety and performance ratings in this study may have triggered participants and raters to consider the dimensions orthogonally, while the single rating of performance used by Rapee and Lim may have been confounded by perceived anxiety.

The results of this study suggest that, in terms of both self and other ratings, nonclinical individuals perform better socially than do individuals with dysthymia, who in turn perform better socially than do individuals with social anxiety disorder. These findings are somewhat contrary to previous findings that observers do not detect differences between socially anxious and nonsocially anxious individuals on ratings of performance (e.g., Rapee & Lim, 1992) or anxiety (Clark & Arkowitz, 1975). The interaction between rating source and diagnosis suggests that socially anxious individuals place a more negative bias on their estimates of their social performance than do either individuals with dysthymia or nonclinical individuals. Thus, it appears that self-ratings of performance by individuals with social anxiety disorder reflect both a kernel of truth and a negative perceptual bias.

While it may indeed be that a negative perceptual bias is inflating self-ratings of anxiety among the clinical groups, with individuals with social anxiety disorder evidencing a greater bias than the participants with dysthymia, it must also be considered that the interaction does not reflect a self-perceptual bias but rather that much of the anxiety experienced is not publicly observable. Indeed, physiological and cognitive manifestations of anxiety would not necessarily be observable by others, and therefore may explain the discrepancy between self and observer anxiety ratings. This explanation seems unlikely, as Alden and Wallace (1995) found a very similar pattern of discrepancy as in this study, when socially anxious participants' ratings of the visibility of their anxiety were compared to observer ratings of anxiety. This alternative hypothesis cannot be discounted, however, and future research should strive to obtain self-ratings of both the perceived visibility of anxiety and degree of anxiety experienced to further elucidate this issue.

It also is arguable that the discrepancy between self and observer ratings of performance and anxiety could have occurred because the observers viewed video recordings of the

interactions, and thus may not have been able to see subtle social nuances occurring during the role-plays. Clearly, future research could rectify this limitation by asking confederates or live observers to provide ratings of anxiety and performance as well. There is evidence, however, that interpersonal factors were not driving the discrepancy. The confederates completed the MAACL-R before and after the unstructured role-play to determine the impact of the interaction on the confederate. If the discrepancies were produced by characteristics of interaction that were perceivable in the role-play but not on the video recording, the discrepancy would have been expected to relate to changes in the MAACL-R subscale scores across the role-play. None of these variables were uniquely related, however, to the self/observer discrepancies in anxiety and performance ratings beyond that already accounted for by percent negative thoughts and, for the anxiety discrepancy, negative evaluation fears.

Finally, the dissimilarities between our findings and those of Rapee and Lim (1992) may be related to subtle differences between the two studies. First, the participants in their study were Australian and ours were from the United States. Thus, the differences observed between the studies may represent sociocultural differences. Secondly, the observers employed by Rapee and Lim were study participants, whereas the observers in the current study were research assistants who were participating in the study. Therefore, the observers used by Rapee and Lira may have made ratings in relation to their own previous or anticipated performance, which could possibly account for the differences between the studies.

Although this study was conducted in a laboratory setting under contrived conditions, there is support for the situational generalizability of the results. The three assessment situations differed in performance difficulty and anxiety induction, but the pattern of the interaction between rating source and diagnosis held consistent across situations. Thus, it is unlikely that the results obtained in this study arose simply as a function of the assessment strategy. The participants, however, may have felt a greater sense of evaluation than they would in a noncontrived social interaction. Indeed, the degree of similarity between social performances in naturalistic and role-played situations remains equivocal (for a review, see Norton & Hope, 2001), although observable anxiety does not appear to amplify if participants are aware they are being assessed (Wessberg, Mariotto, Conger, Farrell, & Conger, 1979).

The inclusion of participants with comorbid anxiety and mood disorders, particularly the social anxiety disorder/dysthymia cross-comorbidity, also enhances the external validity of the results. Indeed, among clinical samples, comorbidity within and across anxiety and mood disorders appears to be the norm, not the exception (Brown & Barlow, 1992; Sanderson et al., 1990). Despite this generalizability advantage, the inclusion of individuals with social anxiety disorder/dysthymia cross-comorbidity also represents a limitation of this study as it blurs the distinction between the clinical groups. When the primary analyses were reanalyzed excluding individuals with cross-group comorbidity, however, estimates of effect size did not change significantly, suggesting that the inclusion of cross-comorbidity did not influence our results.

In addition, the use of only female confederates during the structured and unstructured interactions limits the extent to which we can generalize these results to interactions with

other individuals. Extensions of this study should consider the impact of other confederate-participant pairings such as only men as confederates, same- and opposite-sex pairings, and match or mismatch according to participants' sexual orientation. Indeed, some data suggest confederate gender influences behavior in role-played interactions (e.g., Eisler, Hersen, Miller, & Blanchard, 1975).

Overall, self-ratings of performance by individuals with social anxiety disorder appear to reflect a negative self-perception of one's own performance and a kernel of truth—an accurate awareness that their social performance is somewhat impaired. The clinical control sample demonstrated a similar effect, but of a smaller magnitude, suggesting that such discrepancies may be a key feature of social anxiety disorder. This study also supports previous work that found that although individuals with social anxiety disorder feel much more anxious during social performances than do nonclinical individuals, only a small amount of the anxiety is actually perceived by others.

Our data also provide support for Rapee and Heimberg's (1997) cognitive-behavioral model of social anxiety. Consistent with their model, the results of this study indicate that individuals with social anxiety disorder do not entirely base their self-evaluations on their actual performance and thus may well have a negative mental representation of their functioning during social interactions. Further, this study provides further evidence suggesting that cognitive factors, such as negative evaluation fears and a predominance of negative cognitions, are uniquely related to the degree of negative bias in self-perceptions of one's social performance.

Beyond advancing our understanding of the psychopathology of social anxiety disorder, the findings presented in this paper also hold potential implications for treatment. Notably, the significant discrepancy between self and observer ratings could be used to challenge the beliefs of patients with social anxiety disorder that their anxiety is apparent and obvious to others and their social performance is poor. As well, the results suggest that self-perceptions of performance quality in social situations may reflect an exaggeration of existing performance difficulties. However, the extent to which the performance difficulties experienced by individuals with social anxiety disorder occurred as a function of elevated anxiety or skill deficits is unclear. Practitioners should be cognizant of this uncertainty, and monitor performance quality across treatment to determine if performance varies as a function of anxiety, or if adjunctive skills training may be beneficial.

Acknowledgments – This research was supported in part by grant #MH48751 from the National Institute of Mental Health (NIMH) awarded to the second author. This paper was presented at the annual meeting of the Association for Advancement of Behavior Therapy, New Orleans, 2000.

References

Alden, L. E., & Wallace, S. T. (1995]. Social phobia and social appraisal in successful and unsuccessful social interactions. *Behaviour Research and Therapy, 33,* 497–505.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77–100.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Becker, R. E. (1990). Social skills training for depression. In A. S. Bellack, & M. Hersen (Eds.), *Handbook of comparative treatments for adult disorders* (pp. 88–104). New York: Wiley.

Beidel, D. C., Turner, S. M., Jacob, R. G., & Cooley, M. R. (19891). Assessment of social phobia: Reliability of an impromptu speech task. *Journal of Anxiety Disorders, 3,* 149–158.

Brown, T. A., & Barlow, D. H. (1992). Comorbidity among anxiety disorders: Implications for treatment and DSM-IV. *Journal of Consulting and Clinical Psychology, 60,* 835–844.

Cacioppo, J. Y., Glass, C. R., & Merluzzi, T. V. (1979). Self-statements and self-evaluations: A cognitive-response analysis of heterosocial anxiety. *Cognitive Therapy and Research, 3,* 249–262.

Clark, J., & Arkowitz, H. (1975). Social anxiety and self-evaluation of interpersonal performance. *Psychological Reports, 36,* 211–221.

DiNardo, P. A., & Barlow, D. H. (1988). *Anxiety Disorders Interview Schedule-Revised.* Albany, NY: State University of New York.

Eisler, R. M., Hersen, M., Miller, E. M., & Blanchard, E. B. (1975). Situational determinants of assertive behaviors. *Journal of Consulting and Clinical Psychology, 43,* 330–340.

Glasgow, R. E., & Arkowitz, H. (1975). The behavioral assessment of male and female social competence in dyadic heterosexual interactions. *Behavior Therapy, 6,* 488–498.

Heimberg, R. G., Bruch, M. A., Hope, D. A., & Dombeck, M. (1990). Evaluating the States of Mind model: Comparison to an alternative model and effects of method of cognitive assessment. *Cognitive Therapy and Research, 14,* 543–557.

Heimberg, R. G., Dodge, C. S., Hope, D. A., Kennedy, C. R., & Zollo, L. J. (1990). Cognitive behavioral group treatment for social phobia: Comparison with a credible placebo. *Cognitive Therapy and Research, 14,* 1–23.

Heimberg, R. G., Holt, C. S., Schneier, F. R., Spitzer, R. L., & Liebowitz, M. R. (1993). The issue of subtypes in the diagnosis of social phobia. *Journal of Anxiety Disorders, 7,* 249–269.

Heimberg, R. G., Hope, D. A., Dodge, C. S., & Becker, R. E. (1990). DSM-III-R subtypes of social phobia: Comparison of generalized social phobics and public speaking phobics. *Journal of Nervous and Mental Disease, 178,* 172–179.

Holt, C. S., Heimberg, R. G., Hope, D. A., & Liebowitz, M. R. (1992). Situational domains of social phobia. *Journal of Anxiety Disorders, 6,* 63–77.

Hope, D. A., Heimberg, R. G., & Bruch, M. A. (1995). Dismantling cognitive-behavioral group therapy for social phobia. *Behaviour Research and Therapy, 33,* 637–650.

Hope, D. A., Herbert, J. D., & White, C. (1995). Diagnostic subtype, avoidant personality disorder and the efficacy of cognitive-behavioral group therapy for social phobia. *Cognitive Therapy and Research, 19,* 285–303.

Keppel, G., & Zedeck, S. (1989). *Data analysis for research designs: Analysis of variance and multiple regression/correlation approaches*. New York: Freeman.

Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin, 9,* 371–375.

Leary, M. R., & Kowalski, R. M. (1995). The self-presentation model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 94–112). New York: Guilford.

Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. (1980). Social competence and depression: The role of illusory self-perceptions. *Journal of Abnormal Psychology, 89,* 203–212.

Liebowitz, M. R., Heimberg, R. G., Travers, J., & Stein, M. B. (2000). Social phobia or social anxiety disorder: What's in a name? *Archives of General Psychiatry., 57,* 191–192.

Lubin, B., Zuckerman, M., Hanson, P. G., Armstrong, T., & Rinck, C. M. (1986). Reliability and validity of the Multiple Affect Adjective Check List—Revised. *Journal of Psychopathology and Behavioral Assessment, 8,* 103–117.

Marks, I. M. (1985). Behavioral treatment of social phobia. *Psychopharmacology Bulletin, 21,* 615–618.

Nelson, R. O., Hayes, S. C., Felton, J. L., & Jarrett, R. B. (1985). A comparison of data produced by different behavioral assessment techniques with implications for models of social-skills inadequacy. *Behaviour Research and Therapy, 23,* 1–11.

Norton, P. J., & Hope, D. A. (2001). Analogue observational methods in the assessment of social functioning in adults. *Psychological Assessment, 13,* 59–72.

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy, 35,* 741–756.

Rapee, R. M., & Lim, L. (1992). Discrepancy between self- and observer ratings of performance in social phobics. *Journal of Abnormal Psychology, 101,* 728–731.

Sanderson, W. C., DiNardo, P. A., Rapee, R. M., & Barlow, D. H. (1990). Syndrome comorbidity in patients diagnosed with a DSM-III-R anxiety disorder. *Journal of Abnormal Psychology, 99,* 308–312.

Spitzer, R. L., Williams, J. B., Gibbon, M., & First, M. B. (1989). *Structured Clinical Interview for DSM-III-R.* New York: New York State Psychiatric Institute.

Stopa, L., & Clark, D. M. (1993). Cognitive processes in social phobia. *Behaviour Research and Therapy, 31,* 255–267.

Waiters, K. S., & Hope, D. A. (1998). Analysis of social behavior in individuals with social phobia and nonanxious participants using a psychobiological model. *Behavior Therapy, 29,* 387–407.

Watson, D., & Friend, R. (1969). Measurement of social-evaluation scale. *Journal of Consulting and Clinical Psychology, 33,* 448–457.

Weilage, M., & Hope, D. A. (1999). Self-discrepancy in social phobia and dysthymia. *Cognitive Therapy and Research, 23,* 637–650.

Wenzel, A., & Holt, C. S. (2000). Social phobia and its relationship to clinical syndromes in adulthood. In S. G. Hofmann & E M. DiBartolo (Eds.), *From social anxiety to social phobia: Multiple perspectives* (pp. 130–147). Needham Heights, MA: Allyn & Bacon.

Wessberg, H. W., Mariotto, M. J., Conger, A. J., Farrell, A. D., & Conger, J. C. (1979). The ecological validity of role plays for assessing heterosocial anxiety and skill of male college students. *Journal of Consulting and Clinical Psychology, 47,* 525–535.

Zuckerman, M., & Lubin, B. (1985). *Manual for the MAACL-R: The Multiple Affect Adjective Check List Revised.* San Diego, CA: Educational and Industrial Testing Service.